

Zurek's envariance derivation of Born's rule and measurement

F. Herbut^{1,a}

¹ Serbian Academy of Sciences and Arts, Knez Mihajlova 35, 11000 Belgrade, Serbia

Abstract. Zurek's derivation of Born's rule using envariance (invariance due to entanglement) is considered to capture the probability in full generality, but only as applied to measurement of a quantum observable. Contrariwise, textbook formulations of Born's rule begin with a pure state of a closed, undivided system. The task of this study is to show that a rearrangement of the Zurek approach is possible in which the latter is viewed as giving the probabilities for Schmidt states of an arbitrary composite state vector, and afterwards it is extended to probabilities in a closed, undivided system. This is achieved by determining simultaneously probability and measurement based on the fact that the physical meaning of probability and that of measurement are inextricably dependent on each other.

PACS numbers: 03.65.Ta, 03.65.Ca

1 Introduction

The quantum probability law or Born's rule (as it is also called for historical reasons) is one of the fundamental pillars of physics. All attempts to derive it are of value, and so is Zurek's approach.

This article is focused on the question how one can derive, within the Zurek approach, probabilities for any closed, undivided system. A simple procedure is presented in the next section, and it is pointed out that there is a price to be paid for the simplicity: contradiction with laboratory experience. In section 3 it is shown that this difficulty can be surmounted if one derives also nondemolition measurement of incomplete observables with a finite number of distinct eigenvalues simultaneously with derivation of probability.

Since this study comes somewhere at the end of the Zurek campaign at issue, first a short review of the latter is given.

^a e-mail: fedorh@sanu.ac.rs

Zurek's derivation of Born's rule from entanglement [1], [2] stands out among the similar efforts because entanglement is the generator of decoherence, which, in turn, proved to be a fundamental quantum entity [3], [4], [5], [6], [7]. Zurek made use of entanglement in terms of an entanglement-determined symmetry. He coined the word *envariance* for it. (Initially it was short for environment-assisted invariance, e. g. [8], [2], but later it evolved into the better term: entanglement-assisted invariance [9].)

The initial *envariance* approach [1] gave rise to a number of commentaries and analyses in the literature [10], [11], [12], [13], [14], including that of the present author [15], followed by [16]. Zurek himself commented [2] on the articles of Schlosshauer and Fine [11] and of Barnum [12].

The study of the present author [15] came almost at the end of this series, even after Zurek's second paper [2]. It was aimed at a critical rederivation. It began with a comprehensive theory of twin unitaries - the other face of *envariance*; it was followed by a thorough and critical derivation of Born's rule for Schmidt states (see definition below) from entanglement along the lines of Zurek's reasoning in three stages but with twin unitaries instead of *envariance*. Finally, an attempt was made to extend Zurek's theory to the general case in stages 4 and 5. I think now that this attempt can be done better.

Incidentally, my commentaries on some of the remarks in the preceding articles were left out in the published paper [15]. These remarks gave rise to my attempts at improvement. They can be found in the archive version.

In his latest publication (as far as known to me) [6] Zurek gives on *envariance* only the references [11], [12] and [15] (besides his own). (Perhaps the inventor of the *envariance* derivation does not consider the rest of the articles on the subject as useful contributions.)

My argument in [15] for going beyond Zurek's result was based on ideal measurement and the stipulation that probabilities are predictions for the statistical weights of definite-result sub-ensembles in measurement (cf the passage between relations (22) and (23) *ibid.*). But this same stipulation (cf the last passage in [17]) with ideal measurement was subsequently shown to be by itself sufficient to derive the probability law.

2 Closed, undivided system

A simple derivation is presented in the first subsection. It is critically examined in the second subsection.

2.1 Derivation

The result of Zurek's argument can be shortly put as follows.

Zurek's envariance theorem Let $|\Psi\rangle_{12}$ be an *arbitrary state vector of a composite $(1+2)$ system* in which *the subsystems are entangled*. (The latter requirement is no more than ruling out a tensor product of subsystem state vectors.) Then, if it is written in the form of a *biorthogonal decomposition* (so-called Schmidt decomposition)

$$|\Psi\rangle_{12} = \sum_i \alpha_i |i\rangle_1 |i\rangle_2, \quad \forall i: \alpha_i \in \mathbf{C},$$

$$\forall i, i': \langle i|_1 |i'\rangle_1 = \langle i|_2 |i'\rangle_2 = \delta_{i,i'} \quad (1)$$

(this can always be done, though in general non-uniquely), the probabilities of the *Schmidt states* $\{|i\rangle_1 : \forall i\}$, or, more precisely, of the events (projectors) $\{|i\rangle_1 \langle i|_1 : \forall i\}$, are $\{|\alpha_i|^2 : \forall i\}$ coinciding with the values given by Born's rule.

Remark The result is obviously *symmetric* in the two subsystems. Namely, it is an arbitrary, subjective decision which subsystem we write as first and which as second. The properties of the subsystems displayed in (1) are symmetric under exchange of the latter.

Actually, put in this way, Zurek obtained Born's rule only for Schmidt states. As a justification, entanglement due to measurement-like interaction is invoked suggesting that the environment can be replaced by a measuring instrument or that the latter can be part of the former.

Let us take a different view of Zurek's approach. Let $|\phi\rangle$ be an *arbitrary* state vector of a simple, i. e., undivided, or a composite closed system. ('Closedness' is meant in the sense of lack of quantum correlations with the environment.) Let, further,

$$A = \sum_k a_k |a_k\rangle \langle a_k|, \quad k \neq k' \Rightarrow a_k \neq a_{k'}, \quad (2)$$

be an *arbitrary complete observable* in spectral form. (Completeness of an observable means that no eigenvalue is degenerate.)

We write down the expansion of $|\phi\rangle$ in the eigen-basis of the observable:

$$|\phi\rangle = \sum_k \phi_k |a_k\rangle. \quad (3)$$

To derive probability for any case that is more general than the scope of Zurek's enviance theorem, one must *add* some new requirements to enviance. But the adjoined requirements must fit naturally with probability and entanglement. Since both these concepts are purely mathematical unless one falls back on measurement, we have thus a natural entity to consider along with enviance.

One should have in mind that in any Copenhagen-inspired interpretation of quantum mechanics measurement consists of the dynamical unitary evolution in which the interaction of object and measuring instrument is contained - called premeasurement [18] - and of collapse. In a no-collapse or relative-state approach measurement consists only of premeasurement. All the connection with probability is in premeasurement. Henceforth the terms "measurement" and "premeasurement" are used interchangeably.

Let us take for our premeasurement the simplest one: that of *ideal measurement of a complete observable* A (cf (2)), and let us assume validity of Zurek's enviance theorem. *The first step* in the determination of premeasurement is the requirement of the so-called *calibration condition* (CC) (see my comment on this in the next subsection):

$$\forall k : |a_k\rangle_1 |\chi^0\rangle_2 \rightarrow |a_k\rangle_1 |\chi^k\rangle_2, \quad (4)$$

with $|\chi^0\rangle_2$ as the initial (or ready-to-measure) state of the measuring instrument, and $\{|\chi^k\rangle_2 : \forall k\}$ as the orthonormal pointer states. (All first-subsystem entities appear with the index 1 when dealing with the composite system-plus-instrument system.)

On account of expansion (3) and linearity of the evolution operator, this entails

$$|\phi\rangle_1 |\chi^0\rangle_2 \rightarrow \sum_k \phi_k |a_k\rangle_1 |\chi^k\rangle_2. \quad (5)$$

The final composite state vector of premeasurement appears in (5) in a *biorthogonal* expansion. This is a consequence of the fact that $\{|a_k\rangle_1 : \forall k\}$ and $\{|\chi^k\rangle_2 : \forall k\}$ are orthonormal eigen-vectors of the measured observable and the pointer observable (which is not explicitly made use of) respectively. Hence, Zurek's enviance theorem and the Remark imply that $\{|\phi_k|^2 : \forall k\}$ are the probabilities of the corresponding pointer states $\{|\chi^k\rangle_2 : \forall k\}$.

Measurement applies to individual systems, but also to ensembles of these. The latter fact requires a *second step* in defining measurement, to stipulate the following: *The probabilities of the pointer states $\{|\chi^k\rangle : \forall k\}$ in the final composite state (5) that results from premeasurement, should equal those of the corresponding measured results $\{a_k : \forall k\}$ or $\{|a_k\rangle : \forall k\}$ in the measured state $|\phi\rangle$.* This is the so-called *probability reproducibility condition* (PRC, cf p. 28 in [18]).

It follows that we must *require* the probabilities $\{|\phi_k|^2 : \forall k\}$ to be simultaneously also the sought-for probabilities of the corresponding results.

This simple indispensable requirement achieves the goal to extend the scope of Zurek's envariance theorem to undivided closed systems.

2.2 Critical remarks

Measurement theory assumes knowledge of the probability law. Since we are deriving the latter, we must examine critically the concepts of measurement theory utilized in the above derivation. It goes 'from scratch', by which I mean that we assume the quantum formalism less the probability law and lacking the definition of measurement. One should note that hereby two important concepts are assumed to be valid:

(i) State vectors are taken to represent pure states; more precisely, homogeneous ensembles of equally prepared quantum systems.

(ii) The so-called eigenvalue-eigenstate (e-e) link is assumed. It says that if an observable (Hermitian operator) A , a state vector $|\phi\rangle$, and a real number (eigenvalue) a_k satisfy an eigenvalue relation $A|\phi\rangle = a_k|\phi\rangle$, *then and only then* the observable has the definite value a_k in the state at issue. By 'definite value' one means a non-probabilistic concept: each individual system in the ensemble represented by $|\phi\rangle$ has the value a_k of A .

Naturally, the CC is based on the e-e link. As to PRC, it may have meaning even without a known form of the probability law because it requires only sameness.

Perhaps a serious criticism against the derivation in the preceding subsection is the remark that ideal measurement of a complete observable is actually *impossible* in the physical laboratory.

Such measurement requires a measuring instrument with infinitely many pointer states $\{|\chi^k\rangle : k = 1, 2, \dots, \infty\}$ due to the infinite dimensionality of the state space of a real system. Such an instrument does not exist. In

actuality one can measure only observables with a finite number of distinct eigenvalues, and one does this with the help of a pointer observable with just as many pointer entities.

The project of this study is to show that Zurek's approach allows one to overcome this difficulty.

3 Joint derivation of probability and nondemolition measurement.

We now set the task to derive probability and measurement together from scratch. We take an incomplete observable A , i. e., we allow each of the eigenvalues a_n of A to be arbitrarily degenerate. Then the spectral form in terms of the eigen-projectors $\{P^n : \forall n\}$ of A reads:

$$A = \sum_n a_n P^n, \quad n \neq n' \Rightarrow a_n \neq a_{n'}, \quad (6)$$

and the completeness relation $\sum_n P^n = I$ is valid, where I is the identity operator.

The first necessity for measurement is a measuring instrument provided with a pointer observable B , the eigenvalues b_n and the corresponding eigenprojectors Q^n of which will serve to indicate the results of measurement. One has the spectral decomposition $B = \sum_n b_n Q^n$ with $\sum_n Q^n = I$. (One utilizes the same index n as in (6) because a one-to-one relation between the eigenvalues of the measured observable and the pointer observable must have been established.) The measuring instrument is, further, assumed to be equipped with an initial (or ready-to-measure) state $|\chi^0\rangle$.

The measurement takes place in the form of the evolution (premeasurement)

$$U_{12}(|\phi\rangle_1 |\chi^0\rangle_2) \equiv |\Psi\rangle_{12} = \sum_n \overline{|\Psi\rangle_{12}^n} \quad (7a)$$

where the composite-system term vectors are defined as follows:

$$\forall n : \quad \overline{|\Psi\rangle_{12}^n} \equiv Q_2^n |\Psi\rangle_{12} \quad (7b)$$

(unnormalized vectors are overlined).

The vectors $\overline{|\Psi\rangle_{12}^n}$ are 1-eigen-vectors of Q_2^n and simultaneously the b_n -eigen-vectors of the pointer observable B unless they are zero. (We write shortly Q_2^n instead of $I_1 \otimes Q_2^n$.)

The *probability reproducibility condition* (PRC) was introduced in subsection 2a, and it was pointed out that it is essential for the definition of ensemble measurement. In the present more general case it requires that the probabilities of the pointer events $\{Q^n : \forall n\}$ in the composite state $|\Psi\rangle_{12}$ (cf (7a)) should *equal* those of the corresponding measured results $\{a_n : \forall n\}$ or, equivalently, of $\{P_1^n : \forall n\}$ in the measured state $|\phi\rangle$.

To establish (7a) as premeasurement, we must *require also the calibration condition* (CC) to hold. Let us start with it.

We want to express the claim that if the event P_1^n is certain (in the absolute, i. e., individual-system-in-the-ensemble sense) in an initial state $|\phi\rangle_1$, then so is Q_2^n in the final state $|\Psi\rangle_{12}$ (cf (7a)).

The e-e link allows us to write the initial state in the CC in measurement in the form

$$|\phi\rangle_1 = P_1^n |\phi\rangle_1. \quad (8a)$$

The entire CC then reads: If the initial state of the system is such that (8a) is valid, then the measurement evolution (7a) gives a final state satisfying

$$|\Psi\rangle_{12} = Q_2^n |\Psi\rangle_{12}. \quad (8b)$$

We confine ourselves to *nondemolition* (synonyms: predictive, repeatable, or first kind) measurement in which, by definition, if the measured observable A has a definite value a_n in the initial state $|\phi\rangle$, it is not demolished, i. e., the resulting composite-system state $|\Psi\rangle_{12}$ (cf (7a)) still is an a_n -eigenvalue state of A_1 .

If one has individual systems in mind, then the omitted - demolition, retrodictive, nonrepeatable or second-kind - measurements are not satisfactory, though they are more frequent in the laboratory, because one cannot verify after measurement the just obtained result on the measured individual system for every result.

Incidentally, this investigation bears some resemblance to Zurek's article [19]. But both the details and the purposes differ.

To make the presentation more transparent, we henceforth enumerate those values of the index n for which the term vector $|\Psi\rangle_{12}^n$ in (7a) is non-zero by k .

Thus, in *nondemolition measurement* one joins the following relations to relation (8b)

$$\forall k : \quad |\Psi\rangle_{12}^k = P_1^k |\Psi\rangle_{12}^k. \quad (9)$$

Finally, starting with an arbitrary state vector $|\phi\rangle$, when coupled to the initial state of the measuring instrument and when nondemolition measurement is performed, one obtains

$$U_{12}(|\phi\rangle_1 |\chi^0\rangle_2) = U_{12}\left[\left(\sum_n P_1^n |\phi\rangle_1\right) |\chi^0\rangle_2\right] = \sum_k \overline{|\Psi\rangle_{12}^k} \quad (10a)$$

(the zero terms in the last expression are omitted, and this is emphasized by the change of the enumerating index), so that

$$\forall k : \quad \overline{|\Psi\rangle_{12}^k} = P_1^k \overline{|\Psi\rangle_{12}^k} = Q_2^k \overline{|\Psi\rangle_{12}^k}, \quad (10b)$$

and

$$\forall k : \quad \overline{|\Psi\rangle_{12}^k} = \langle\phi|_1 P_1^k |\phi\rangle_1^{1/2} |\Psi\rangle_{12}^k, \quad (10c)$$

where $|\Psi\rangle_{12}^k$ is the corresponding unit vector.

Relation (10c) is due to the fact that the unitary evolution does not change the norm of any term $P_1^n |\phi\rangle_1 |\chi^0\rangle_2$ in (10a). (Now we see that the omitted zero terms are those for which $\langle\phi|_1 P_1^n |\phi\rangle_1 = \|P_1^n |\phi\rangle_1\|^2 = 0$, i. e., those n -values for which P_1^n takes $|\phi\rangle_1$ into zero.)

The terms $\overline{|\Psi\rangle_{12}^k}$ in the final state in (10a) are biorthogonal due to (10b). When one writes down any Schmidt decomposition of each state vector $|\Psi\rangle_{12}^k = \sum_l \phi_{kl} |kl\rangle_1 |kl\rangle_2$, and one makes use of (10c), one arrives at

$$|\Psi\rangle_{12} = \sum_k \overline{|\Psi\rangle_{12}^k} = \sum_k \langle\phi|_1 P_1^k |\phi\rangle_1^{1/2} \sum_l \phi_{kl} |kl\rangle_1 |kl\rangle_2. \quad (11)$$

Here one has $\langle kl|_1 |\bar{k}\bar{l}\rangle_1 = \delta_{k,\bar{k}} \delta_{l,\bar{l}} = \langle kl|_2 |\bar{k}\bar{l}\rangle_2$, i. e., one obtains a Schmidt decomposition of the entire composite final state $|\Psi\rangle_{12} = \sum_k \overline{|\Psi\rangle_{12}^k}$.

Now *Zurek's envariance theorem* (with the Remark) are applicable. They say that the sought-for probability of each Schmidt state $|kl\rangle_2$ in the final state $|\Psi\rangle_{12} = \sum_{k'} \overline{|\Psi\rangle_{12}^{k'}}$ is given by the expression $(\langle\phi|_1 P_1^k |\phi\rangle_1) |\phi_{kl}|^2$.

To be able to avail ourselves of the PRC, we need to evaluate the probability of Q_2^k in $\sum_{k'} \overline{|\Psi\rangle_{12}^{k'}}$ for each k value. In order to use the obtained probability values of the Schmidt states $|kl\rangle_2$, we need to realize that

$$Q_2^k \sum_l |kl\rangle_2 \langle kl|_2 = \sum_l |kl\rangle_2 \langle kl|_2,$$

or symbolically $(\sum_l |kl\rangle_2 \langle kl|_2) \leq Q_2^k$ (a sub-projector relation; proof is given in Appendix A). Hence, there exists a possibly nonzero projector $(Q_2^k)'$ that is orthogonal to the sub-projector, and such that $Q_2^k =$

$\left(\sum_l |kl\rangle_2 \langle kl|_2 \right) + (Q_2^k)'$. But $\left(\sum_l |kl\rangle_2 \langle kl|_2 \right)$, just like Q_2^k , takes $\sum_{k'} \overline{|\Psi\rangle_{12}^{k'}}$ into the corresponding term $\overline{|\Psi\rangle_{12}^k}$ implying that $(Q_2^k)'$, even if it is nonzero, acts as zero on $\sum_{k'} \overline{|\Psi\rangle_{12}^{k'}}$.

At this stage it seems unavoidable to introduce the *additivity assumption*: If a projector is an orthogonal sum of projectors $P = \sum_m P_m$, then the probability $p(P)$ of the corresponding event is the *sum* of the probabilities of the term events $p(P) = \sum_m p(P_m)$ in any state.

Since we need to allow our event (projector) P_1^k to have an infinite-dimensional range, we must require the equality to hold true even if the projector sum has an infinite number of non-zero terms. This is called σ -*additivity* (it implies additivity).

I can offer only intuitive plausibility support for the σ -additivity stipulation. It starts with pointing to the fact that every event has an opposite event (every projector has its ortho-complementary event). But, in order to avoid perhaps unnecessary repetition, I ask the reader to read my detailed explanation in subsection V.E. of my former article [15] on the envariance derivation.

On account of Zurek's envariance theorem and σ -additivity, we obtain that the probability of the event (projector) $\left(\sum_l |kl\rangle_2 \langle kl|_2 \right)$, in $\sum_{k'} \overline{|\Psi\rangle_{12}^{k'}}$ is $\langle \phi|_1 P_1^k |\phi\rangle_1$ as seen utilizing (11) and the probabilities of the Schmidt states $|kl\rangle_2$ in (11).

The probability of Q_2^k in $\sum_{k'} \overline{|\Psi\rangle_{12}^{k'}}$ is, on account of additivity, that of $\left(\sum_l |kl\rangle_2 \langle kl|_2 \right)$ plus that of $(Q_2^k)'$, which is zero. Hence also the probability of Q_2^k is in $\sum_{k'} \overline{|\Psi\rangle_{12}^{k'}}$ equal to $\langle \phi|_1 P_1^k |\phi\rangle_1$.

Finally, invoking the PRC, we obtain $\langle \phi|_1 P_1^k |\phi\rangle_1$ as the probability of the eigen-event P_1^k in the state $|\phi\rangle_1$. Thus, *Born's rule for an arbitrary state vector and an arbitrary event* is derived.

Further generalization is straightforward. For the reader's convenience, it is presented in Appendix B.

In view of the fact that every real (in the sense of laboratory-realizable) measuring instrument must have a *finite* number of distinct pointer events $\{Q_2^n\}$, one can make the following comment. If at least one of the eigenvalues a_n of the measured observable A is infinitely degenerate, then, in spite of the infinite dimensionality of the state space, the index n may take a finite number of values. This is why we are dealing with measurement that is, in principle, realizable in the laboratory.

4 Concluding remarks

In the above, perhaps cumbersome, but essentially straightforward intertwined derivation of measurement and probability we have seen that for deriving the former, the CC and the PRC seemed to be unavoidable as crucial partial determinations of the measurement concept. (Measurement becomes completely determined when so is probability.)

The e-e link and σ -additivity were inescapable for the project of deriving probability.

In every derivation of probability one is worried about possible circularity: could some assumption(s) be equivalent to the entire or to a part of the probability law?!

Let me turn to a deep critical observation in the article of Schlosshauer and Fine [11]. In their Concluding remarks they say:

”...a fundamental statement about any probabilistic theory: We cannot derive probabilities from a theory that does not already contain some probabilistic concept; at some stage, we need to put probabilities in to get probabilities out.”

At this place we must recall Gleason’s theorem [20], which says that a σ -additivity observing probability function on the entire lattice of events is equivalent to a general quantum state (and *vice versa*) via the trace rule $\text{tr}(E\rho)$, where E is an arbitrary projector and ρ is an arbitrary density matrix. Gleason’s fundamental theorem makes σ -additivity the reason for circularity in the sense of the above quote.

One should not hold this insight against Zurek’s approach because it is not a competitor to Gleason’s theorem; it is a complementary effort.

The derivation of the probability law from entanglement and additional assumptions serves the purpose to shed light on the ‘miracle’ how this law comes about. One does understand, at least mathematically, how superposition of tensor-product (hence uncorrelated) state vectors gives rise to entanglement. Zurek’s enviance derivation project shows, at least to a large extent, how this brings about the trace rule (equivalent to Born’s rule) via its enviance symmetry.

Finally, let me point out that Zurek, who launched the enviance project to which this study attempts to be a modest contribution, is going on pre-

sending new ideas within the project [21].

ACKNOWLEDGEMENT The author is indebted to prof. Arthur Fine for calling his attention to the need to clarify that collapse plays no role in the argument of this article. I am also grateful to an unknown referee, whose thorough analysis and competent, helpful criticism enlightened me and contributed substantially to making the final text more readable. Naturally, for the remaining weaknesses I am the only one to blame.

Appendix A

Lemma Let $|\Psi\rangle_{12} = \sum_l \alpha_l |l\rangle_1 |l\rangle_2$ be a Schmidt decomposition of a bipartite state vector, let $\rho_2 \equiv \text{tr}_1(|\Psi\rangle_{12}\langle\Psi|_{12})$ be the reduced density operator of the second subsystem, and let Q_2 be a second-subsystem projector such that $Q_2 |\Psi\rangle_{12} = |\Psi\rangle_{12}$. Then $Q_2 \sum_l |l\rangle_2 \langle l|_2 = \sum_l |l\rangle_2 \langle l|_2$, symbolically $(\sum_l |l\rangle_2 \langle l|_2) \leq Q_2$.

Proof $Q_2 \rho_2 = \text{tr}_1(Q_2 |\Psi\rangle_{12}\langle\Psi|_{12}) = \rho_2$. As well known, any Schmidt decomposition is accompanied by $\rho_2 = \sum_l r_l |l\rangle_2 \langle l|_2$, where $r_l (\equiv |\alpha_l|^2)$ are the positive eigenvalues and $|l\rangle_2$ are the corresponding eigenvectors of ρ_2 . (We assume that the above Schmidt decomposition exhibits only non-zero terms.) Hence $\sum_{l'} r_{l'} Q_2 |l'\rangle_2 \langle l'|_2 = \sum_{l''} r_{l''} |l''\rangle_2 \langle l''|_2$. Taking the $\langle l|_2 \dots |l\rangle_2$ matrix element, one obtains $\langle l|_2 Q_2 |l\rangle_2 = 1$. From this $Q_2 |l\rangle_2 = |l\rangle_2$ and the claimed relation ensue. \square

Appendix B

Completion from pure state to general state.

Let us rewrite the obtained expression of the probability in a form suitable for further generalization. Let P be an arbitrary event (projector). Then, we have derived that its probability in a pure state $|\phi\rangle$ is

$$p(P, |\phi\rangle) = \langle\phi| P |\phi\rangle = \text{tr}(P |\phi\rangle \langle\phi|) \quad (B.1)$$

Let ρ be a density matrix representing a **proper mixed state**. Let, further,

$$\rho = \sum_k w_k |\phi\rangle_k \langle\phi|_k \quad (B.2)$$

be one way of writing the state as a mixture of pure states. The physical meaning of (B.2) is that ρ , as an ensemble, can be prepared by mixing

the pure states in (B.2) in the proportions given by the statistical weights w_k . More precisely, the numbers N_k of the individual systems making up the pure sub-ensembles $|\phi\rangle_k\langle\phi|_k$ relate to each other in the same way as the w_k do. Hence,

$$\forall k : w_k = N_k / N, \quad N \equiv \sum_{k'} N_{k'}. \quad (B.3)$$

Let us imagine that P is measured on each individual system in the ensemble ρ , and that the event occurs on \bar{N}_k systems in the k -th sub-ensemble, altogether on $\sum_k \bar{N}_k$ systems in the entire ensemble. Since the probability applies to a random individual system, the sought for probability is $\sum_k \bar{N}_k / N$ (cf (B.3)). Thus,

$$p(P, \rho) = \sum_k \bar{N}_k / N = \sum_k \left(N_k / N \right) \left(\bar{N}_k / N_k \right) = \sum_k w_k \text{tr}(P |\phi\rangle\langle\phi|) = \text{tr}(P\rho). \quad (B.4)$$

Evidently, all the previous relations (B.1)-(B.3) have been utilized.

If ρ is a density matrix representing an **improper mixed state**, then, by definition, there exists another system with which our system is entangled, and together they are in a pure state. Denoting our system by index 1, the other system by index 2, and the state of the $(1+2)$ system by $|\Psi\rangle_{12}$, the improper mixed state is determined by the partial trace

$$\rho_1 = \text{tr}_2 \left(|\Psi\rangle_{12} \langle\Psi|_{12} \right). \quad (B.5)$$

Making use of (B.4) and (B.5), it is straightforward to evaluate the probability $p(P_1, \rho_1)$ of an arbitrary first-subsystem event P_1 .

$$p(P_1, \rho_1) = \text{tr}_{12} \left(P_1 (|\Psi\rangle_{12} \langle\Psi|_{12}) \right) = \text{tr}_1 [P_1 (\text{tr}_2 (|\Psi\rangle_{12} \langle\Psi|_{12}))] = \text{tr}_1 (P_1 \rho_1). \quad (B.6)$$

(Note that indices are used unnecessarily also for full traces for transparency.)

We have obtained the result that the formalism does not distinguish proper and improper mixed states.

Naturally, all claims made in this Appendix are standard reasoning. They are required to complete the derivation.

References

- [1] W. H. Zurek, Phys. Rev. Lett. **90**, 120404 (2003); quant-ph/0211037.
- [2] W. H. Zurek, Phys. Rev. **A 71**, 052105 (2005); quant-ph/0405161).
- [3] E. Joos, H. D. Zeh, C. Kiefer, D. Giulini, J. Kupsch, and I-O. Stamatescu, *Decoherence and the Appearance of a Classical World in Quantum Theory* (Berlin, Springer, 2003).
- [4] H. D. Zeh, <http://www.rzuser.uni-heidelberg.de/~as3/index.html>
- [5] W. H. Zurek, Rev. Mod. Phys. **75**, 715 (2003); quant-ph/0105127.
- [6] W. H. Zurek, in *Many Worlds?*, edited by S. Saunders, J. Barrett, A. Kent, and D. Wallace (Oxford, Oxford University Press, 2010) pp. 409-432.
- [7] J. Bub *Interpreting the Quantum World* (Cambridge, England, Cambridge University Press, 1997) subsection 8.1.
- [8] W. H. Zurek in *Science and Ultimate Reality: From Quantum to Cosmos*, edited by J. D. Barrow, P. C. W. Davies, and C. H. Harper (Cambridge, England, Cambridge University Press, 2004) pp. 121-137.
- [9] W. H. Zurek, Nature Phys. **5**, 181 (2009).
- [10] C. M. Caves, <http://info.phys.unm.edu/caves/reports/ZurekBornderivation.pdf>
- [11] M. Schlosshauer, and A. Fine, Found. Phys. **35**, 197 (2005); quant-ph/0312058v3.
- [12] H. Barnum, quant-ph/0312150 (2003).
- [13] U. Mohrhoff, Int. J. Quant. Inf. **2**, 221 (2004); quant-ph/0401180.
- [14] Ph. Blanchard and E. Brüning, Phys. Lett. **A 355**, 180 (2006).
- [15] F. Herbut, J. Phys. A: Math. and Theor. **40**, 49 (2007); quant-ph/0611220.
- [16] Ph. Blanchard, Eur. Phys. J. **159**, 37 (2008).
- [17] F. Herbut, J. Phys. A: Math. and Theor. **40**, 10549 (2007); arXiv:0708.1932.

- [18] P. Busch, P. K. Lahti, and P. Mittelstaedt *The Quantum Theory of Measurement*, Report Series (Turku, Finland, Department of Physics, University of Turku, 1991).
- [19] W. H. Zurek, Phys. Rev. **A 76**, 052110 (2007).
- [20] A. M. Gleason, J. Math. Mech. **6**, 885 (1957).
- [21] W. H. Zurek, Phys. Rev. Lett. **106**, 250402 (2011).